DOCUMENT RESUME

ED 138 647 TM 006 306

AUTHOR Thissen, David M.
TITLE Response Latency and Individual Differences: An Integration in a General Latent Trait Model.
PUB DATE [Apr 77]
NOTE 22p.; Paper presented at the Annual Meeting of the American Educational Research Association (61st, New York, New York, April 4-8, 1977)

EDRS PRICE MF-$0.83 PC-$1.67 Plus Postage.
DESCRIPTORS Cognitive Ability; Cognitive Tests; *Individual Differences; Item Analysis; *Mathematical Models; Probability; *Reaction Time; *Response Style (Tests); Statistical Analysis
IDENTIFIERS *Latent Trait Models

ABSTRACT

Few phenomena in psychology have been studied more than reaction time. However, while response latency now provides the basic data for research in many areas of cognitive psychology, little work has been done on the problem of incorporating response latency in the measurement of individual differences. A solution to this problem requires the construction of a mathematical model relating reaction time to individual and test item variation. The present research is concerned with the development and empirical validation of such a model. The model developed here provides a simultaneous analysis of test item responses and the latencies associated with those responses. The method is sufficiently general to be used in the analysis of a wide variety of data, ranging from responses to conventional ability tests to data collected in experimental problem-solving situations. The application of this model to a set of spatial visualizing ability test data is discussed. The results indicate that the model provides a parsimonious summary of the raw data, and estimates of individual differences useful in further analysis. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED138647

Response Latency and Individual Differences: An Integration in a General Latent Trait Model

David M. Thissen

The University of Kansas

Paper presented at the 1977 meeting of the American Educational Research Association, April 5, New York City.

Session 4.01

TM006 306

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Abstract

Few phenomena in psychology have been studied more than reaction time. However, while response latency now provides the basic data for research in many areas of cognitive psychology, little work has been done on the problem of incorporating response latency in the measurement of individual differences. A solution to this problem requires the construction of a mathematical model relating reaction time to individual and test item variation. The present research is concerned with the development and empirical validation of such a model. The model developed here provides a simultaneous analysis of test item responses and the latencies associated with those responses. The method is sufficiently general to be used in the analysis of a wide variety of data, ranging from responses to conventional ability tests to data collected in experimental problem-solving situations. The application of this model to a set of spatial visualizing ability test data is discussed. The results indicate that the model provides a parsimonious summary of the raw data, and estimates of individual differences useful in further analysis.

Response Latency and Individual Differences: An Integration in a General Latent Trait Model

It has been claimed that there are two kinds of experimental cognitive psychologists: the "percent correct" kind and the "reaction time" kind. Some cognitive psychologists have attempted to use both response variables, but the different sorts of error models required for the statistical analyses have made that difficult. On the other hand, psychologists concerned with individual differences in cognitive ability have, for the most part, used test item responses in a "percent correct" framework. In the individual differences paradigm, psychometric models, called "latent trait" models, have been developed for the analysis of a variety of forms of item response data, using statistical methods appropriate for the nominal or graded format of the raw data (see Rasch, 1960; Lord & Novick, 1968; Samejima, 1969, 1973; and Bock, 1972 for illustrations of the development of such models). These models yield estimates of individual ability and of item attributes (such as difficulty and discrimination) which are on an interval scale with approximately Gaussian error of estimation and calculable error variance.

The purpose of the present paper is the development of a latent trait model for paired item-response, response-time data which uses both the item responses and the latencies to estimate parameters of individual and item performance. The goal of this line of research is more precise ability estimation and item analysis, through the

use of more information (the latencies). The result of this effort may be viewed both as a useful procedure in its own right, as a test-scoring procedure for individual measurement and item analysis, and as a first stage in a more complex program of data analysis. In the latter case, the latent trait model is used as a mechanism to reduce the highly multivariate item response data to summary parameters for each person and item. Since the parameters, estimated by maximum likelihood, have approximately Gaussian error of estimation, standard least squares techniques may be used to test hypotheses about sources of inter-individual or inter-item variation using the summary parameters as data. The original data need not be assumed to be complete; parameter estimates are available in the presence of a substantial amount of missing data. And the error of measurement of the raw data is not assumed to be Gaussian, which is an attractive feature. Correct/incorrect responses are more appropriately treated as arising from a binomial distribution, and latency distributions are frequently skewed; lognormal error will be assumed for the observed latencies. Latent trait models do assume that the component structure of the task involved in the test items is homogenous; that is, that the data arise from a so-called "one-factor" test. So this is not a technique for examining the "stage-structure" of cognitive response mechanisms (as in, for example, Sternberg, 1969). It is a technique for examining between-individual or between-item differences with respect to a homogenous response mechanism.

A latent trait model specifies the probability of a categorical response (like correct/incorrect) or the probability density for a

continuous response variable (like latency) as a function of individual and item parameters. The following description will be of one of a class of latent trait models which simultaneously describe categorical item responses and latency.

Denote the response to a binary item r_{ij} if the response of person i to item j is correct; the associated latency is t_{ij}. The Birnbaum (1968) logistic model for the probability of a correct response is

$$P(r_{ij}) = \psi\{a_j(\theta_i + c_j)\}, \tag{1}$$

where

$$\psi(z) = 1/\{1 + \exp(-z)\}$$

is the logistic function. θ_i is the "effective ability" of individual i. (The term "effective ability" is borrowed from White (1973) to distinguish this ability estimate, which includes whatever effect on ability accrues from allowing the examinee all the time he wants, from more conventional ability estimates from speeded or time-limited tests.) The term c_j is the easiness of item j. Both θ_i and c_j are measured on a z-scale. The item discrimination parameter, a_j, reflects the item-trait correlation. Figure 1 shows two item characteristic curves, which are plots of the function (1), probability of correct response against ability. Both items in Figure 1 have the same easiness, c; the steep curve is the item of high discrimination, while the other curve is one of lower discrimination. The shape of the curves, from the function $\psi(z)$, is extremely similar to the normal ogive used is scaling and early test theory.

Insert Figure 1 about here

The model for latency is

(2) $\log(t_{ij}) = \mu + s_i + u_j - b\{a_j(\theta_i + c_j)\} + \varepsilon_{ij}$,

in which $\varepsilon_{ij} \sim N(0,\sigma^2)$. This is a regression-type model, in which one of the terms, $\{a_j(\theta_i + c_j)\}$, is the combination of effective ability and easiness used to predict correct response probability. This term, weighted by b (which converts its units to logseconds), reduces the predicted response latency for high ability and/or high easiness, and increases time for negative ability or easiness. The general structure of the model was proposed by Furneaux (1960), but could not be conveniently tested statistically at that time. The extent of the influence on latency of the ability-easiness term is determined by a_j, the item-trait regression. The other terms of the latency model summarize effects on latency not correlated with the latent trait; μ is an overall mean response log-time, and s_i and u_j are individual and item slowness parameters which reflect processing time associated with individuals and items, but not related to the probability of a correct response. The expression (2) gives predicted response latency densities like those in Figure 2 for three values of s, holding all of the other parameters constant.

Insert Figure 2 about here

This model has been applied to a set of test data collected by R. Darrell Bock, using his own test of spatial visualizing ability, the "Bock-Fitzgerald" test. The responses are those of sixty sub-

urban London high school students. The Bock-Fitzgerald test has eighteen items like the one in Figure 3 (top) in which the examinee's task is to pick the two circles at the bottom which are planar rotations of the figure at the top; both responses must be correct to be scored correct, and latency is measured to the second response. The remaining fifteen items of the test are like the one in the lower half of Figure 3; they are similar to the preceding parts, except that the top circle is omitted and the only cues available are the lines which overlap the edge.

Insert Figure 3 about here

The model was fitted by joint maximum likelihood, using an iterative gradient estimation scheme. The goodness-of-fit of the functional forms included in the model may be tested in a variety of ways. Figure 4 shows the fit of the standardized deviations of log latency from the model; the deviations have been divided into fractiles and the histobars hung from the theoretical curve so that discrepancies between model and data are shown at the bottom, as deviations around the straight line. A χ^2 statistic for the goodness-of-fit of the proportions in these ten fractiles to the assumed normal distribution is 5.8 on 7 degrees of freedom, which is non-significant. The data do not noticeably violate the assumption of normal error in the log metric. Similar item goodness-of-fit χ^2s are available for the approximation of the proportion of correct responses across

ability to the logistic curve; for the entire test, the total of these χ^2s is 157.8 on 128 degrees of freedom, which is also non-significant. While these tests do not establish that the model is "correct," they do establish that the error variation is approximately as predicted by the model, and not incompatible with it.

Insert Figure 4 about here

Some of the results of this application are illustrated in Figure 5, which displays slowness plotted against effective ability. The first outstanding feature of the plot is that the two variables are highly correlated, $r = .82$. The parameters have managed to reflect the obvious: that as slowness (carefulness?) increases, effective ability, which controls probability correct, increases. The interesting point comes in the analysis for sex differences. A sex difference is usually found on speeded spatial tests; males have higher mean ability. In this case, neither slowness nor effective ability exhibits a significant sex difference. However, bivariate analysis of variance yields a significant sex difference ($F = 3.58$, $p < .05$). Discriminant analysis reveals that the dimension of difference is the contrast between effective ability and slowness; in other words, propensity for correct response per unit time. Thus, this analysis explicitly finds the fact which has been the basis for speeded testing of spatial ability for the last forty years: What we have called individual differences in spatial ability are

actually individual differences in rate of processing. Of course, under speeded testing conditions, or with problems so difficult that they take so long that other factors, like memory decay, have an effect, this difference in rate of processing may become a difference in absolute ability to do the problems. But such absolute differences in probability correct may only be highly indirect expressions of the individual differences in processing. This view of spatial ability is compatible with the findings of Shepard and his colleages (Shepard & Metzler, 1971; Metzler & Shepard, 1974; Shepard & Judd, 1976), that mental rotation, which is the core of a spatial test, is a process with a well-defined rate within individuals.

Insert Figure 5 about here

This latent trait analysis of timed testing was quite successful with the spatial test data, supplying information unavailable with standard tests of spatial ability. Models similar to that described here have been tried with other spatial data, experimental word-perception data, and data from the Matching Familiar Figures test (Kagan, 1964) with similar success. These general models allow an investigator to simultaneously analyze variation in correct response probability and latency across individuals or items (or experimental stimuli), without concern for reasonable amounts of

missing data and without violating the assumptions of strong statistical analysis. Models such as these should prove to be an important step in allowing the use of the most appropriate response format in cognitive testing, which will allow us to increase the precision, and hopefully the validity, of psychological measurement.

References

Birnbaum, A. Some latent trait models and their use in inferring an examinee's ability. In F. Lord & M. Novick (Eds.), Statistical theories of mental test scores. New York: Addison-Wesley, 1968.

Bock, R. D. Estimating item parameters and latent ability when responses are scored in two or more nominal categories. Psychometrika, 1972, 37, 24-51.

Furneaux, W. D. Intellectual abilities and problem-solving behavior. In H. J. Eysenck (Ed.), The handbook of abnormal psychology. London: Pitman, 1961.

Kagan, J., Rosman, B.L., Day, D., Albert, J., & Phillips, W. Information processing in the child: Significance of analytic and reflective attitudes. Psychological Monographs, 1964, 78, (1, Whole No. 578).

Lord, F., & Novick, M. R. Statistical theories of mental test scores. Reading, Mass.: Addison-Wesley, 1968.

Metzler, J., & Shepard, R. N. Transformational studies of the internal representations of three-dimensional objects. In R. Solso (Ed.), Theories of cognitive psychology. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1974.

Rasch, G. Probabilistic models for some intelligence and attainment tests. Copenhagen: Nielson and Lydiche (for Danmarks Paedagogiske Institut), 1960.

Samejima, F. Estimation of latent ability using a response pattern of graded scores. Psychometrika Monograph, No. 17, 1969.

Samejima, G. Homogenous case of the continuous response model. Psychometrika, 1973, 38, 203-219.

Shepard, R. N., & Judd, S. A. Percpetual illusion of rotation of three-dimensional objects. Science, 1976, 191, 952-954.

Shepard, R. N., and Metzler, J. Mental rotation of three-dimensional objects. Science, 1971, 171, 701-703.

Sternberg, S. Memory scanning: Mental process revealed by reaction time experiments. American Scientist, 1969, 57, 421-437.

White, P. O. Individual differences in speed, accuracy, and persistance. In H.J. Eysenck (Ed.), The measurement of intelligence. Lancaster, England: Medical and Technical Publishing Co., Ltd., 1973.

Figure Caption

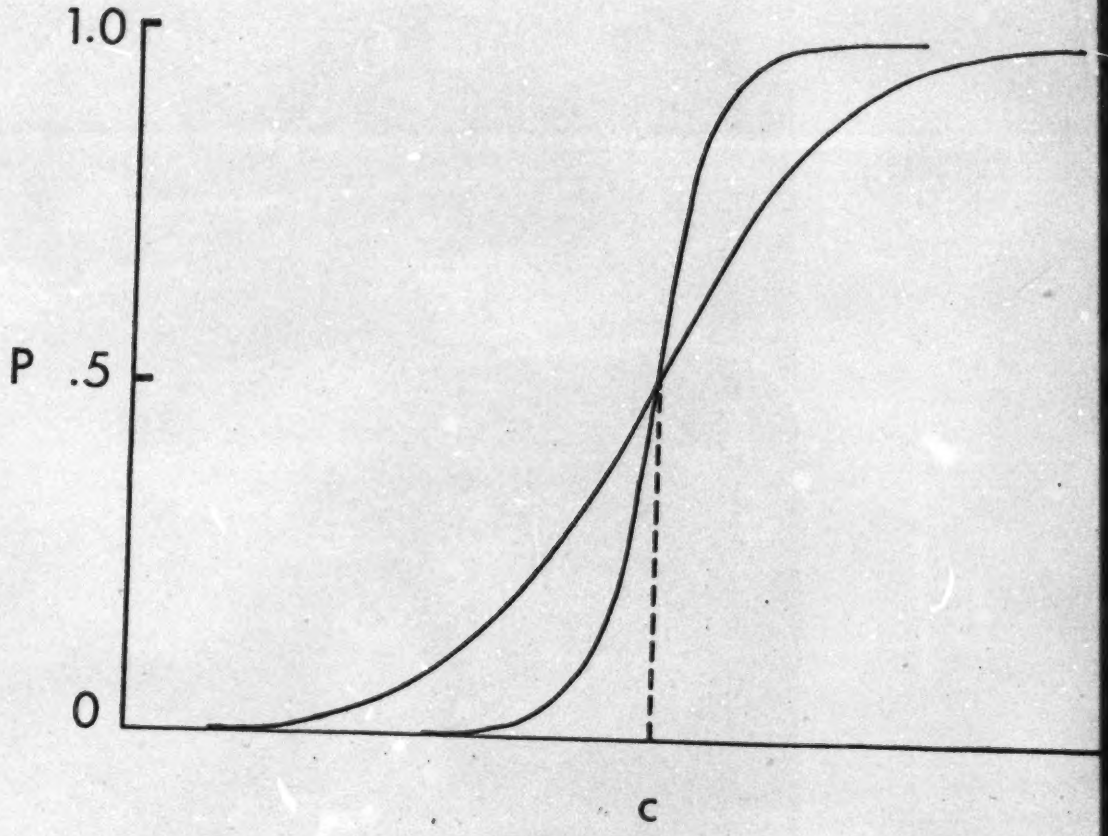
Figure 1. Item characteristic curves for two items with different discriminations; probability correct against effective ability.

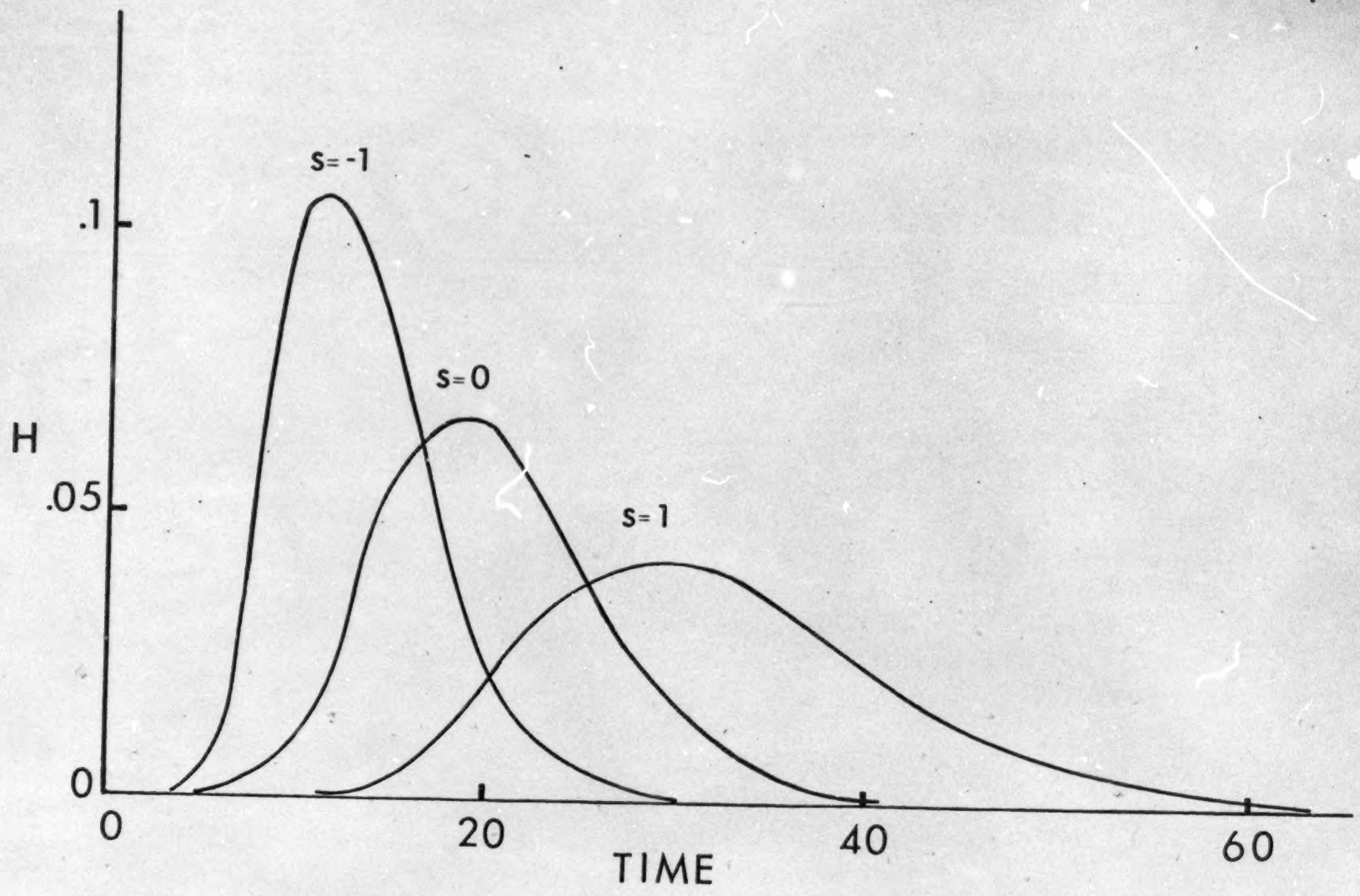
Figure 2. Distributions of response latency for different values of s; ability and item characteristic held constant.

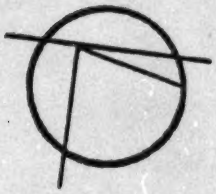
Figure 3. Bock-Fitzgerald test items.

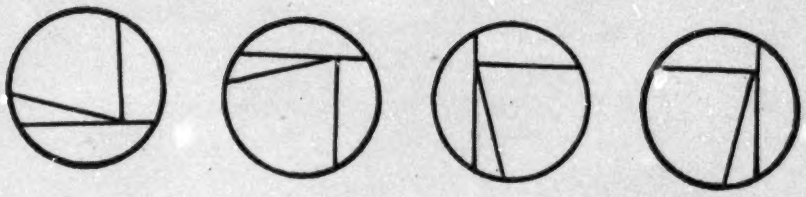
Figure 4. Hanging Histogram of the standardized deviations of log latency from the model.

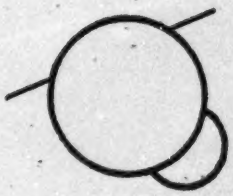
Figure 5. Effective ability and slowness plotted for the males and females in the London sample (open circles females, closed squares males).

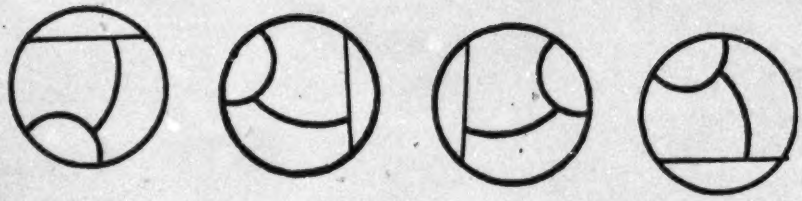

1.0
P .5
0
c

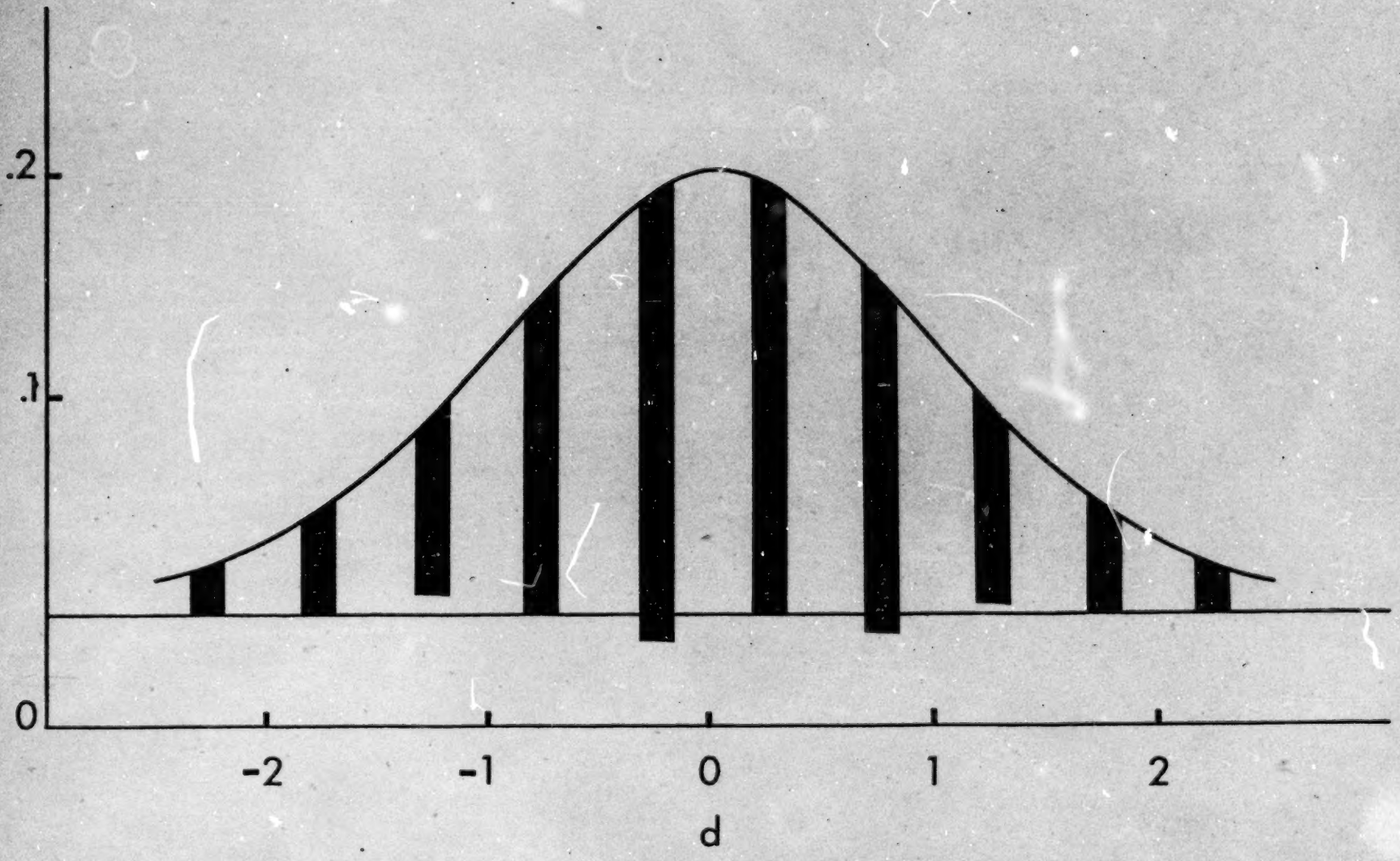

s=-1
s=0
s=1
.1
.05
0
H
0
20
40
60
TIME

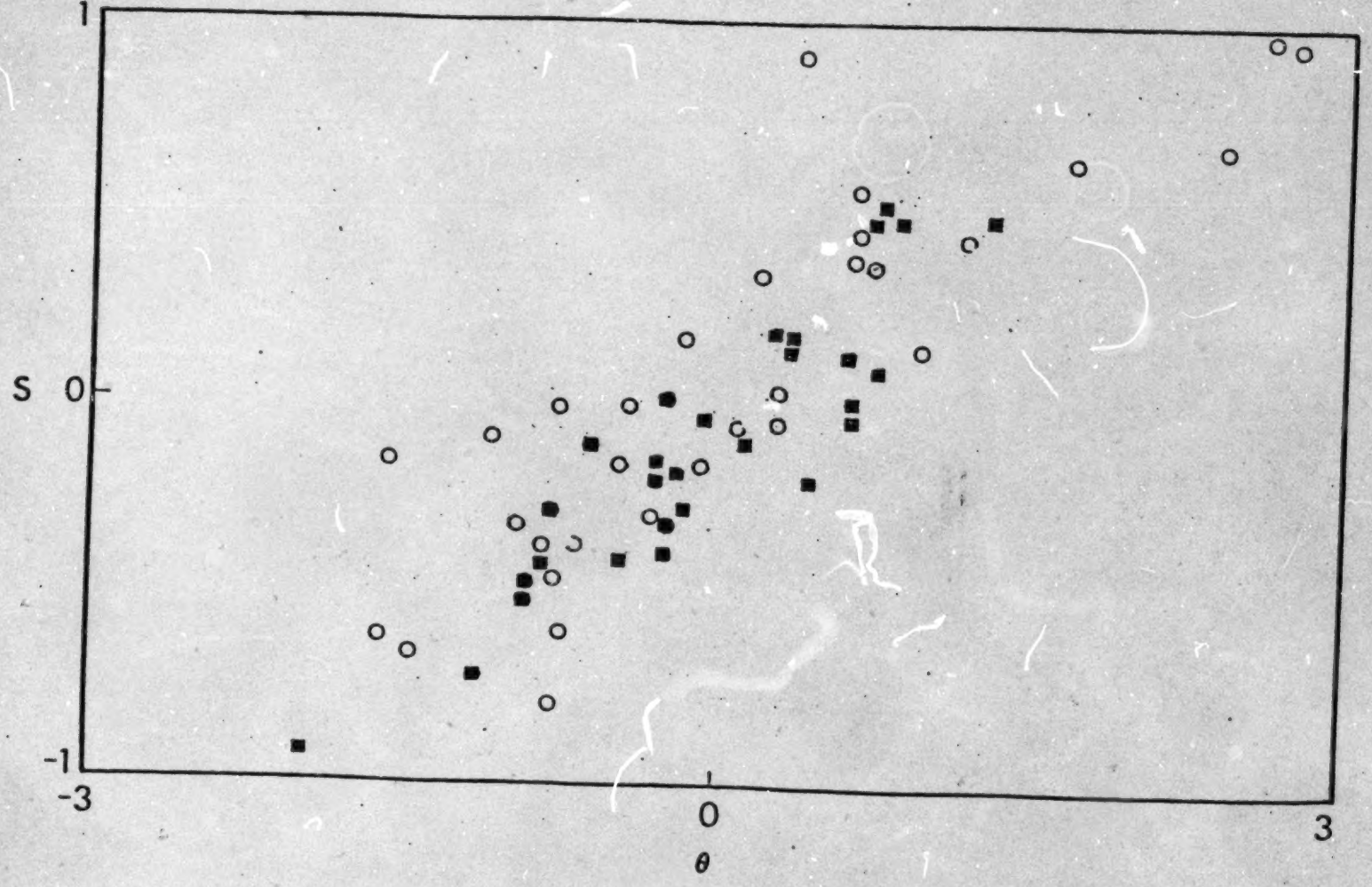









.2
.1
0
-2
-1
0
1
2
d


1
S 0
-1
-3
0
3
θ